

# Pennsylvania County Voting Anomaly Analysis

S. Stanley Young, PhD, FASA, FAAAS: *Revision 11-8-20*

This report looks at Pennsylvania county voting, 2008 to 2020. One data set has 67 rows, with one row for each county. The first few rows are given here.

RowID	PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden 2020	Dif	Rank	Democratic	Republican	Other	Total	%Dem
1	Adams	17633	15091	14219	17919	2,271.3	20	19090	36862	10824	66776	28.6
2	Allegheny	373153	352687	367617	396767	32,281.3	3	555649	263952	138999	958600	58.0
3	Armstrong	11138	9045	7178	7130	-1,990.3	60	13903	22848	4848	41599	33.4
4	Beaver	40499	37055	32531	37389	694.0	24	53417	41701	14063	109181	48.9

This report is in the form of text describing a question of interest, a figure and discussion of the figure.

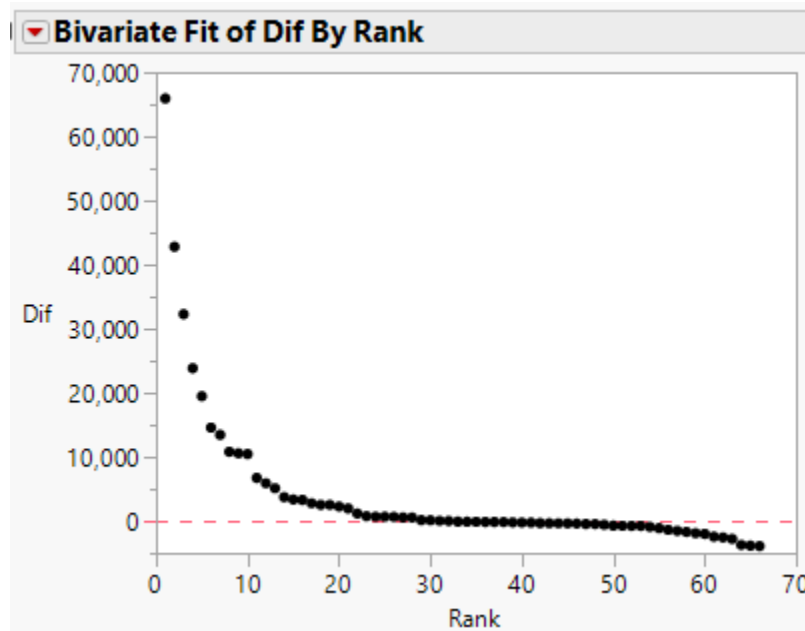
## Item 1 —

In the current data set 8Oct2020, several counties have noteworthy characteristics. Both counties are heavily Democratic, 58% and 76%. Allegheny County has votes for Biden 32,281 above the average vote counts for Obama 2008, Obama 2012 and Hillary 2016. Curiously, Philadelphia ~31,000 votes below the average of the previous three presidential elections. Both counties are unusual.

RowID	PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden N	Dif 2	Rank2	Democratic	Republican	Other	Total	%Dem
2	Allegheny	373153	352687	367617	415737	51,251.3	2	555649	263952	138999	958600	58.0
51	Philadelphia	595980	588806	584025	558264	-31,339.7	67	815218	117919	128938	1062...	76.8

## Item 2 —

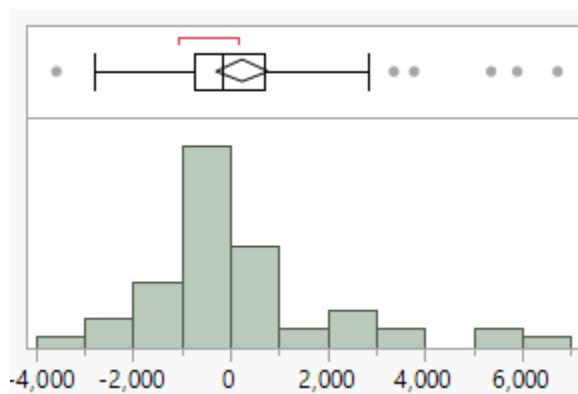
Are there particular counties where voting does **not** follow previous voting patterns? We computed the difference in Biden 2020 versus the average of Obama 2008, Obama 2012 and Hillary 2016. The name of this variable is Dif. We plot Dif versus the rank of the difference. The largest Dif gets rank 1; the 2<sup>nd</sup> largest Dif gets rank 2, etc.



We see that for most counties there is very little difference in this election relative to previous elections. In some counties the Biden count is lower than previous Democratic counts. The same data can be displayed as a histogram, Item 3.

### Item 3a —

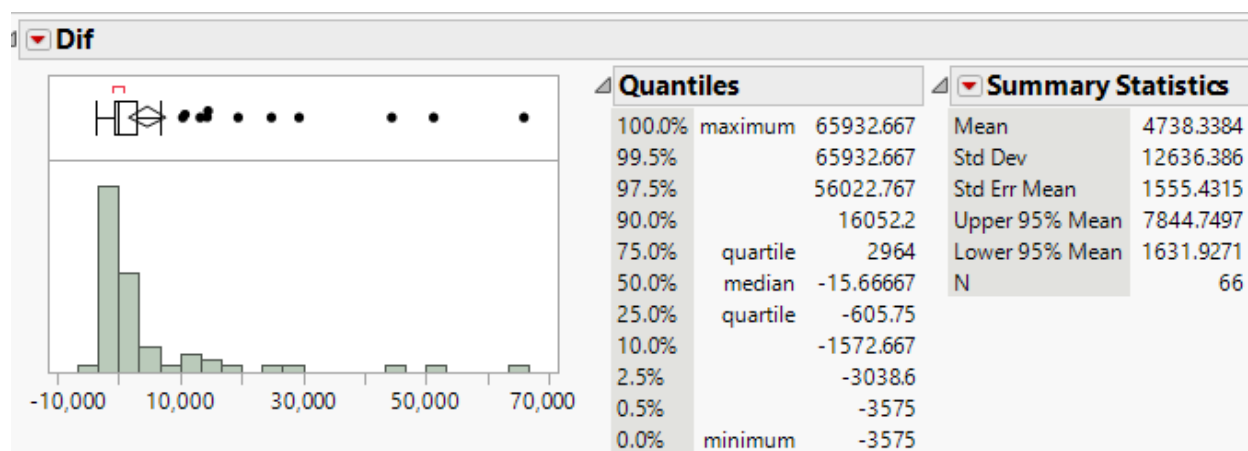
We examine the bulk of the data, omitting for now those counties with a large increase in voting.



We expect little change in vote totals versus the average of previous votes. For most of the counties that appears to be true. In fact, there are more negative Dif values; note the large bar just below 0. Many counties, 33, showed little enthusiasm for Biden as the Biden counts are negative (below the average of previous elections).

### Item 3b —

We now look at the histogram for all the counties, omitting Philadelphia.



On the left of the figure, from -5,000 to about 6,000 we see bars that resemble a normal distribution; See Item 3a. The values above 10,000 appear to be outliers. An outlier is an unusual number relative to other numbers in the collection. It is unusual to see a gain of 10,000 votes or more; reexamine Item 2.

### Item 4a —

See the outlier counties on the next page. These are the increase of votes Biden got over the average of the prior three presidential elections.

Consider Montgomery County. Obama/Hillary vote counts ranged from 233,000 to 256,000. Biden received 313,000. The eleven outlier counties together provide over **299,000** excess votes.

PA Counties	Dif	Rank	%Dem
Montgomery	65,932.7	1	48.8
Allegheny	51,251.3	2	58.0
Chester	44,456.3	3	40.3
Bucks	29,380.3	4	43.0
Delaware	24,889.7	5	47.7
Lancaster	19,482.7	6	32.6
Cumberland	14,582.0	7	33.9
Northampt...	14,434.0	8	45.7
Lehigh	13,307.0	9	48.4
Dauphin	10,838.3	10	45.3
York	10,471.7	11	33.3

#### Item 4b —

The changes in vote counts from Hillary 2016 to Obama 2012 were mostly negative. Many of the same counties come up having large positive changes, esp. Allegheny, Chester, and Montgomery. Note: In 2020 there were five counties with a higher difference than was seen in 2016/2012. Five counties has more votes added than the max number of votes added in 2012/2016.

PA Counties	Hillary - Obama 2012
Montgomery	22726
Chester	17371
Allegheny	14930
Bucks	6539
Delaware	5610
Lehigh	3041
Centre	2912
Cumberland	2718
Lancaster	2612
Union	71
Butler	34

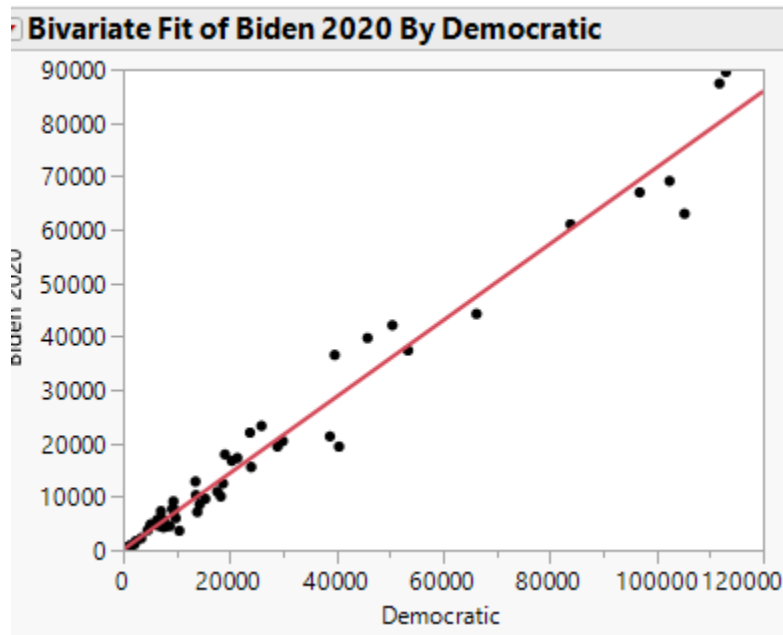
#### Item 5 —

We examine the Biden votes against the number of Democratic registered voters. In all these counties, esp. Montgomery and Chester counties, there are more Biden votes than registered Democrats... Note: It is not statistically reasonable to get more Democratic votes, by ~38,000, than there are registered Democratic voters.

PA Counties	Dif 3
Montgomery	38588
Chester	36642
Delaware	10209
Lancaster	5774
Bucks	2479
Cumberland	1512

### Item 6 —

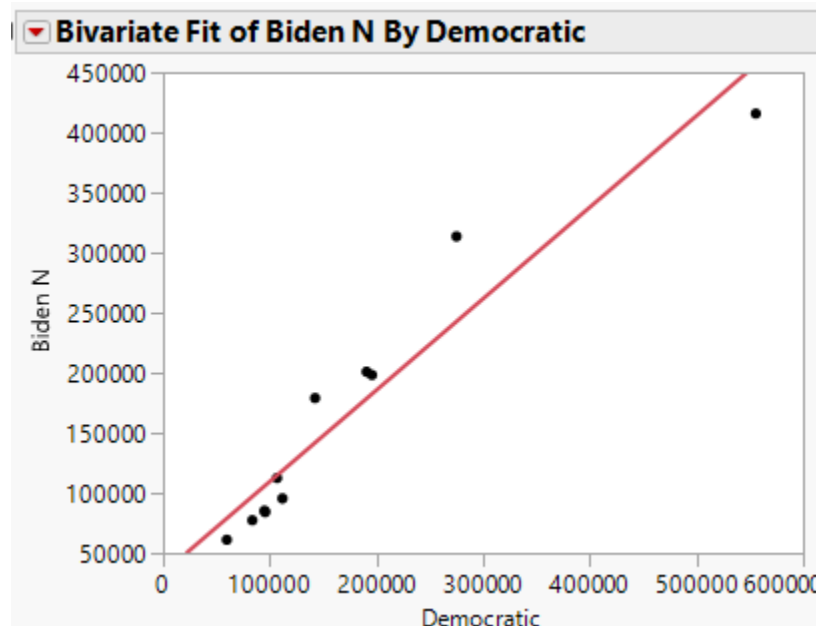
We seek to estimate the fraction of registered Democratic voters that voted for Biden. We want an unbiased estimate so the 11 outlier counties plus Philadelphia were removed from the analysis. Fifty-five (55) counties were used for simple linear regression.



The data are fit well with a simple line.  $\text{Biden 2020} = 111.10562 + 0.7150147 * \text{Democratic}$   
This shows that in the “normal” counties, 70%± of registered Democratic voters vote.

### Item 7 —

We seek to estimate the fraction of registered Democratic voters that voted for Biden among the 10 outlier counties. (We want an unbiased estimate, so we removed Allegheny County as so many Democrat voters appear not to have voted.) As a result, ten (10) counties were used for simple linear regression.



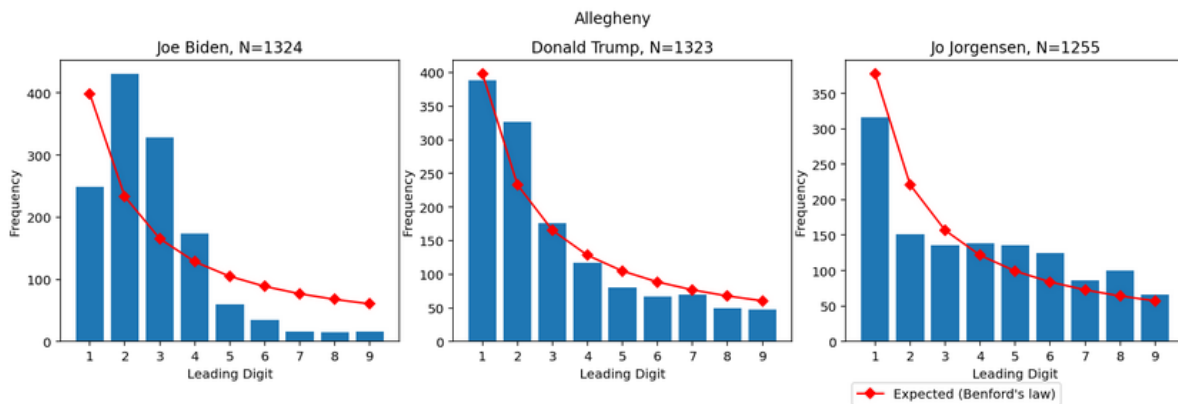
The data are somewhat reasonable fit with a simple line.  $\text{Biden N} = 33552.252 + 0.7607002 * \text{Democratic}$ .

The conclusion is that in these ten suspect counties, 76% of PA registered Democratic voters supposedly voted. We have something of a contradiction. In counties that delivered increase votes for Biden a higher proportion of registered Democrat voters voted than otherwise in the state (which was 70%±).

### Item 8 —

Allegheny County is quite odd. One technique to detect fraud from numerical data is to use the so called Benford's Law\*. That law asserts that the number 1 should appear more frequently as the lead digit in a precinct voter count. The number 2 should be less frequent, etc. The distribution, with no fraud, should decline exponentially. That method has been applied to Allegheny County's voting precincts. See [cjph8914/2020\\_benfords](https://www.cjph8914/2020_benfords).

#### Allegheny, PA



The shape of the voter count first digits follows Benford's Law rather closely for Trump, but does not for Biden. This result lends support for the county analysis that Allegheny County is anomalous.

### SUMMARY:

1. Philadelphia and Allegheny Counties are unusual. They have high Democratic registration. Allegheny precinct votes fail the Benford Law\* test (a standard voter fraud method) whereas Trump does not
2. The votes for Biden are unusually high for eleven counties, reporting over 299,000 votes in excess of expectations. These deviations are legitimate reasons to insist on closely monitored recounts.
3. 70%± of registered PA Democrat voters vote in a national election. After elimination of Allegheny County, 76% of PA registered Democrat voters voted in the nine suspect counties. That is a unexpected statistical aberration.
4. Another way to track down fraudulent votes is to look closely at how many of the votes did little or no down-ticket voting. When manufacturing votes, it is too time consuming to vote for other office holders.

\*"[The Theory and Applications of Benford's Law](#)" (Steven J. Miller, editor)